

Columbia University

ECBM E4040 Neural Networks and Deep Learning. Fall 2020.

Project

Team ID: YEAH

Group Member: Jinzhu Yang(jy3024), Jiayun Ni(jn2722), Saier Gong(sg3772)

In this project, we aim to reconstruct the Residual Attention Network(Attention-56 and Attention-92) and Naive Attention Model(NAL) and reproduce the error rates by training them with CIFAR-10, CIFAR-100 and ImageNet datasets, referencing the paper of Residual Attention Network for Image Classification.

In this notebook, we demonstrate the name of each jupyter notebook that works on different models and the loss and accuracy of it. Some tables of summarized outputs would also be given to demonstrate the results more straightforward and clearly.

NaiveAttention56_CIFAR10_sg.ipynb

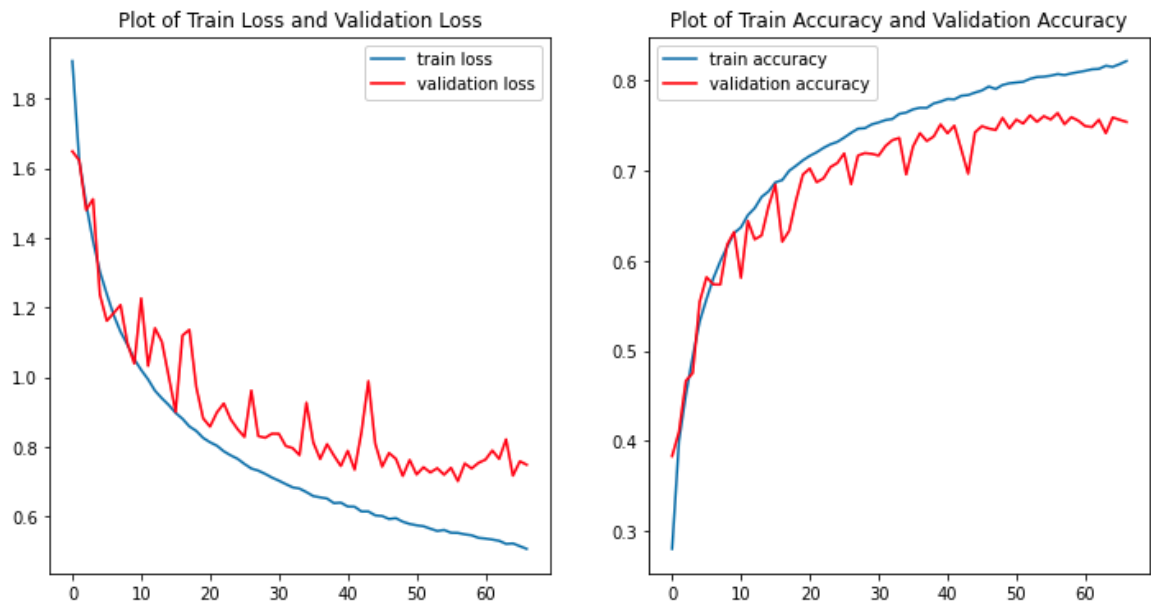
This jupyter notebook works on the model of naive attention learning with the architecture of attention56 model using CIFAR-10 dataset. The test accuracy is 0.8205.



NaiveAttention92_CIFAR10_jn.ipynb

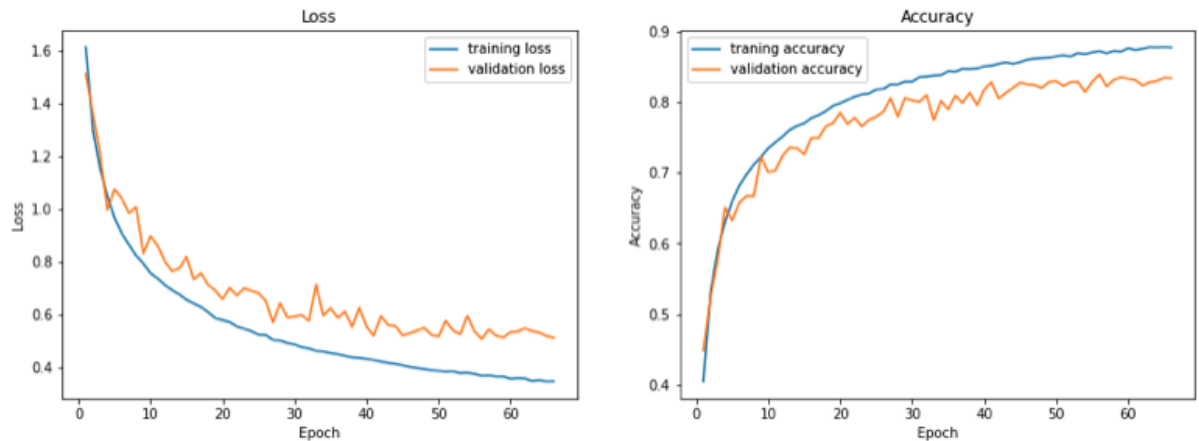
This jupyter notebook works on the model of naive attention learning with the architecture of attention92 model using CIFAR-10 dataset. The test accuracy is 0.7443.

Plot of Accuracy and Loss



ResidualAttention56_CIFAR10_sg.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention56 model using CIFAR-10 dataset. The test accuracy is 0.8341.



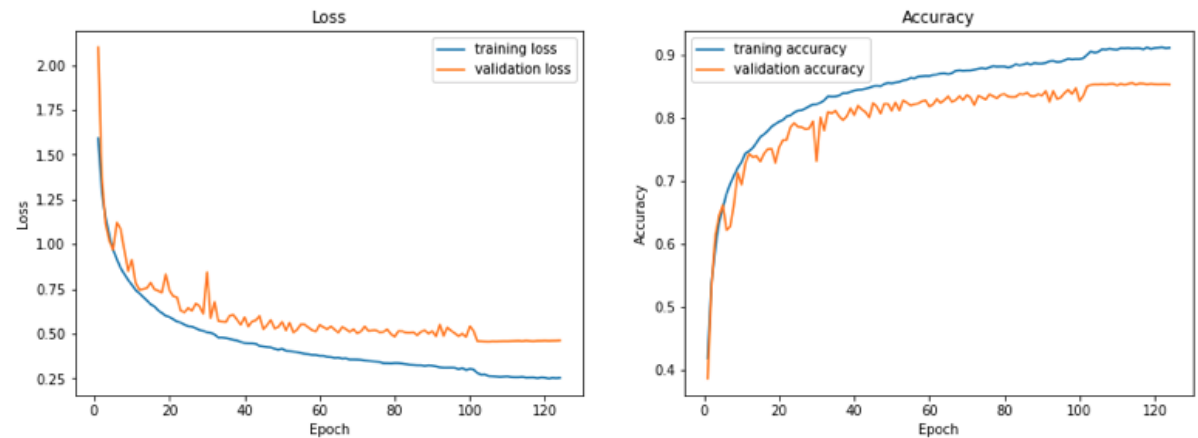
ResidualAttention56_Channel_CIFAR10_sg.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention56 model with channel attention method using CIFAR-10 dataset. The test accuracy is 0.8084.



ResidualAttention56_Spatial_CIFAR10_sg.ipynb

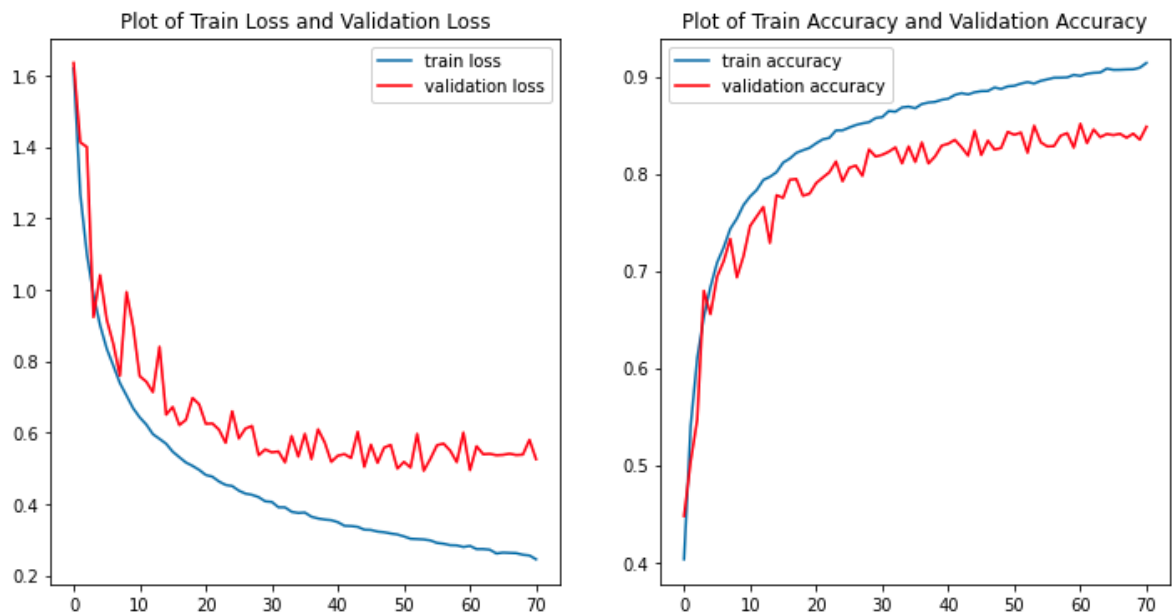
This jupyter notebook works on the model of residual attention network with the architecture of attention56 model with spatial attention method using CIFAR-10 dataset. The test accuracy is 0.8084.



ResidualAttention92_CIFAR10_jn.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention92 model using CIFAR-10 dataset. The test accuracy is 0.8379.

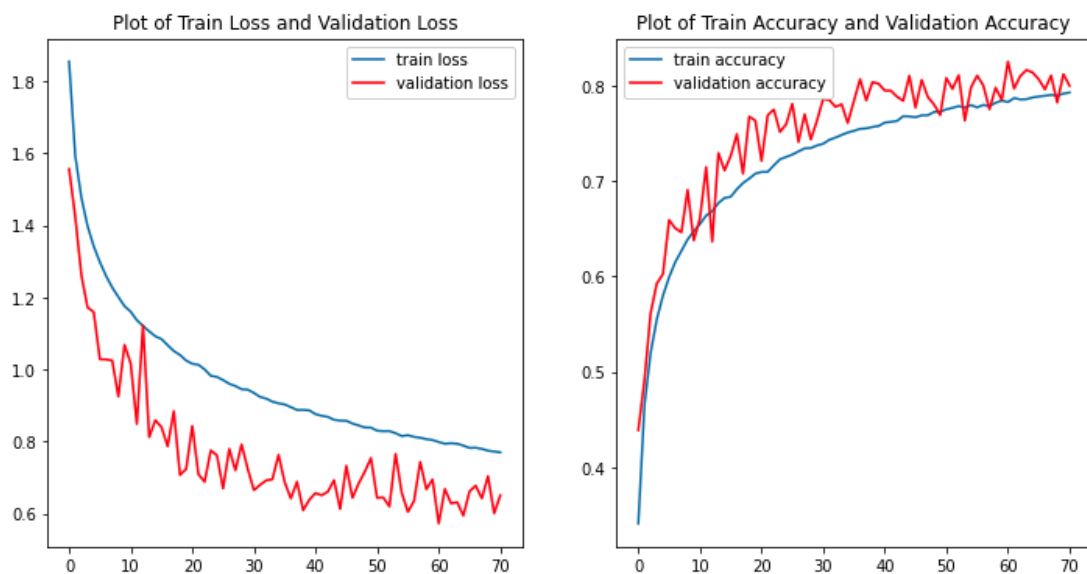
Plot of Accuracy and Loss



ResidualAttention92_CIFAR10_Noise10_jn.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention92 model with noise level of 10% using CIFAR-10 dataset. The test accuracy is 0.7936.

Plot of Accuracy and Loss



ResidualAttention92_CIFAR10_Noise30_jn.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention92 model with noise level of 30% using CIFAR-10 dataset. The test accuracy is 0.7501.

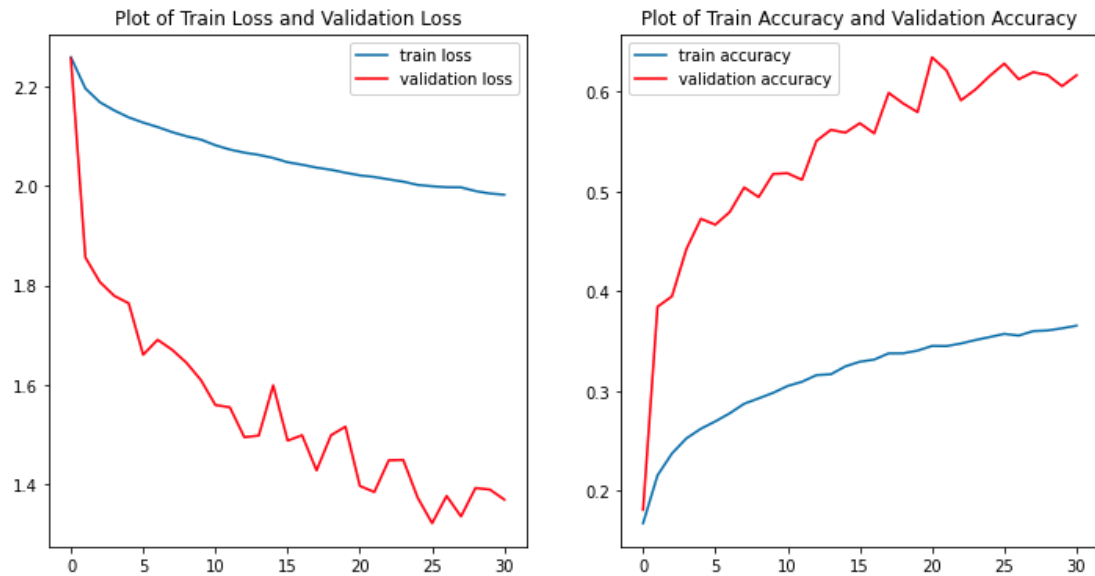
Plot of Accuracy and Loss



ResidualAttention92_CIFAR10_Noise50_jn.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention92 model with noise level of 50% using CIFAR-10 dataset. The test accuracy is 0.6126.

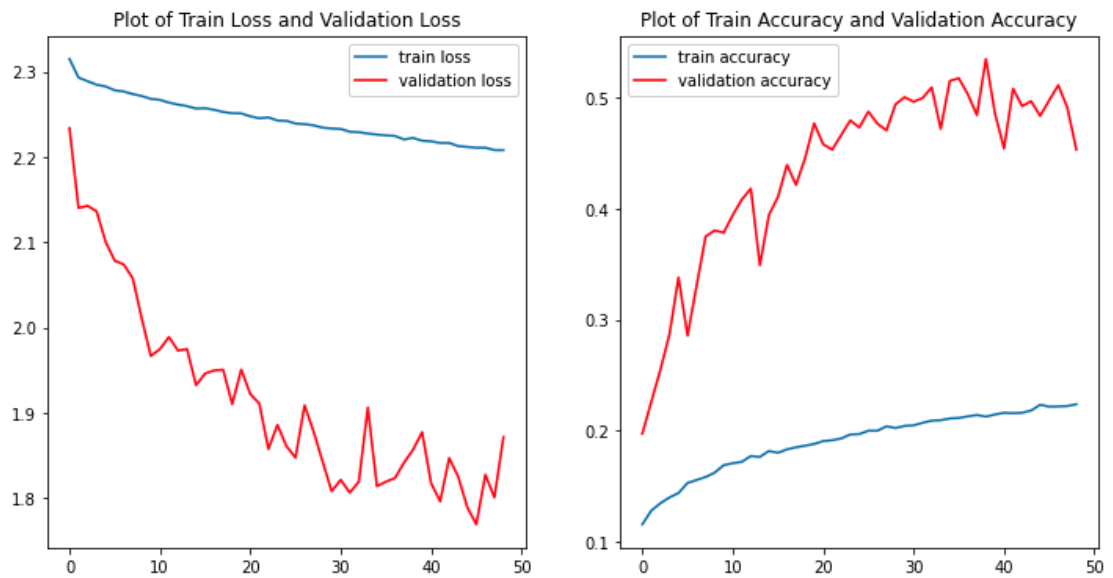
Plot of Accuracy and Loss



ResidualAttention92_CIFAR10_Noise70_jn.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention92 model with noise level of 70 using CIFAR-10 dataset. The test accuracy is 0.4517.

Plot of Accuracy and Loss



ResidualAttention56_CIFAR100_sg.ipynb

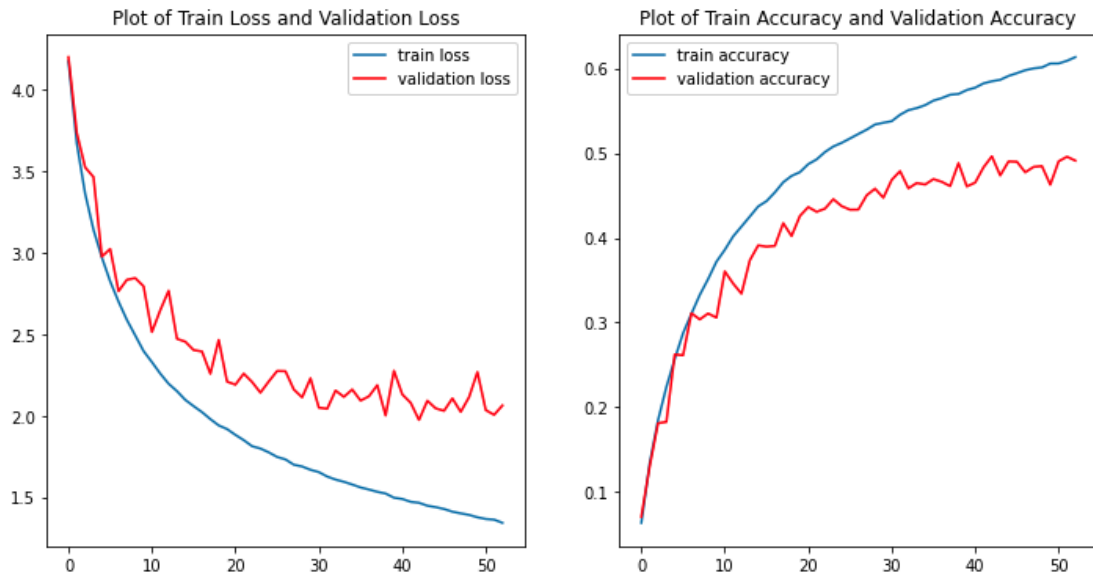
This jupyter notebook works on the model of residual attention learning with the architecture of attention56 model using CIFAR-100 dataset. The test accuracy is 0.5298.



ResidualAttention92_CIFAR100_jn.ipynb

This jupyter notebook works on the model of residual attention network with the architecture of attention92 model using CIFAR-10 dataset. The test accuracy is 0.5013.

Plot of Accuracy and Loss



main_ImageNet.ipynb

This jupyter notebook works on the model of residual attention learning with the architecture of attention56 model using ImageNet dataset. The top-1 error rate is 0.5875 and the top-5 error rate is 0.019.

Summary

Tables below show the same result as above but in the format of tables that would demonstrate the results more clearly and informatively with comparisons among models built.

This table show the error rate and training time of the Attention Residual Learning model and Naive Residual Learning model built using attention56 and attention92 and trained by CIFAR-10 dataset.

Network	ARL(Top-1 err. %)(Training Time)	NAL(Top-1 err. %)(Training Time)
Attention-56	16.59(2h25min8s)	17.95(2h59min58s)
Attention-92	16.21(2h40min24s)	25.57(2h26min3s)

This table shows the error rate and training time of the Residual Attention Network built using different types of attention modules and trained by CIFAR-10 dataset.

Activation Function	Attention Type	Top-1 err. (%)	Training Time
$f_1(x)$	Mixed Attention	16.59	2h25min8s
$f_2(x)$	Channel Attention	19.16	1h56min53s
$f_3(x)$	Spatial Attention	15.36	4h43min24s

This table demonstrates the error rate and training time of the Residual Attention Network built using attention92 model with different level of noise and trained by CIFAR-10 dataset.

Noise Level	Attention-92 err.(%)	Training Time
10%	20.64	2h40min55s
30%	24.99	1h50min55s
50%	38.74	1h20min20s
70%	54.83	1h51min42s

This table demonstrates the error rate and training time of the Residual Attention Network built using attention56 and attention92 model and trained by CIFAR-10 and CIFAR-100 datasets.

Network	CIFAR-10(Training Time)	CIFAR-100(Training Time)
Attention-56	16.59(2h25min8s)	47.20(3h18min)
Attention-92	16.21(2h40min24s)	49.87(1h57min50s)

This table listed the training accuracy, validation accuracy, top-1 test error rate and top-5 test error rate of the Residual Attention Network built using attention56 and trained by ImageNet dataset.

Network	Training Accuracy	Validation Accuracy	Test Top-1 Error	Test Top-5
Attention-56	0.3614	0.3605	0.5875	0.001

